DOCUMENT RESUME

ED 346 157 TM 018 505

AUTHOR Arter, Judith A.; Stiggins, Richard J.
TITLE Performance Assessment in Education.
PUB DATE Apr 92
NOTE 12p.; Paper presented at the Annual Meeting of the American Educational Research Association (San Francisco, CA, April 20-24, 1992).
PUB TYPE Information Analyses (070) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Academic Achievement; *Accountability; Curriculum Development; *Educational Assessment; Elementary Secondary Education; Evaluation Problems; Higher Education; Measurement Techniques; *Outcomes of Education; *Student Evaluation; Teacher Education; *Teacher Evaluation; Teacher Role; Test Construction
IDENTIFIERS *Performance Based Evaluation; Teaching to the Test

ABSTRACT

The combination of increasing demands for accountability and the desire to measure a variety of complex educational outcomes makes the use of performance assessment an essential addition to the tools used to profile student achievement. Performance assessment in education is valuable for student assessment and for the assessment of teacher and principal performance. Lessons learned to date in the practice of performance assessment include: (1) the need for clear targets; (2) the need for an array of assessment tools; (3) the need for training; (4) cost, time, and technical issues; and (5) issues associated with high-stakes testing, such as restricting curriculum, teaching to the test, and other negative effects. Educators find themselves in a dilemma, caught between the complex outcomes that require performance assessment and the prohibitive costs of such assessments. The instructional usefulness of performance assessment is currently limited, but making teachers partners in the assessment process improves the quality of performance assessment and its instructional usefulness. A major unresolved performance assessment problem is that teachers do not have the training to use all that performance assessment offers. The nation does not appear to have the resources to solve this national problem. There is a 36-item list of references. One table summarizes applications of performance assessment in education. (SLD)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☑ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

JUDITH A. ARTER

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

ED346157

PERFORMANCE ASSESSMENT IN EDUCATION

Judith A. Arter
Richard J. Stiggins

Northwest Regional Educational Laboratory
101 S.W. Main, Suite 500
Portland, OR 97204
(503)275-9500

Paper presented as part of the symposium *Performance Assessment Methods: What Can We Learn From Research In Other Professions*, national meeting of the American Educational Research Association, San Francisco, April 1992.

TM018505

BEST COPY AVAILABLE

PERFORMANCE ASSESSMENT IN EDUCATION

Introduction

Performance assessment is certainly not new to education. Teachers have always been observers of student behavior (Stiggins & Bridgeford, 1985). Oral examinations have been used to determine student progress and learning for hundreds if not thousands of years. And beyond these practical applications, performance assessment has been the focus of scholarly activity for decades. Each of the past four decades has produced at least one major update of research and development in performance assessment (Ryans & Frederick, 1951; Glaser and Klaus, 1962; Fitzpatrick and Morrison, 1977; and, Berk, 1986).

What is new in the 1990s is the "holy Grail" emphasis currently being attached to performance assessment in education. The combination of increasing demands for accountability and the desire to measure for a wide variety of more complex educational outcomes makes the use of performance assessment in all its guises--portfolios, assessment centers, systematic classroom observation, structured tasks--an essential addition to the array of tools used to profile student achievement.

Because of the headlong rush to use performance assessment methodology, we have gained a great deal of experience over the past few years both with its appropriate and inappropriate use. For this reason, this is a excellent time to think about educational performance assessment in all its forms and to ponder what we've learned so far.

Overview of Performance Assessment in Education

Performance assessments are used in many ways in education because of the multitude of purposes for these assessments, achievement targets to be addressed, and populations to assess, including teachers, students, and administrators. Table 1 presents a summary of such applications. We will briefly explain these applications and illustrate each with a few examples.

Insert Table 1 about here

Assessment of Student Performance

Purposes. Assessments of student performance serve a variety of purposes. As classroom tools, they can inform specific instructional decisions made by teachers, students and parents, and serve as teaching and learning tools for both teachers and students. Teachers use performance assessment methodology to engage students in the assessment of their own and each other's performance as a means of becoming more accomplished performers.

In fact, one of the most important developments in the classroom use of this methodology over the past decade has been the realization that the entire process of performance assessment can be a powerful instructional tool. For example, consider the six-trait analytical scoring method used for student writing in grades 3-12 in Oregon. This procedure was originally developed in Beaverton School District in 1984 by teachers seeking to improve upon holistic scoring as a means of providing feedback to student writers. The six traits of ideas, organization, voice, word choice, sentence fluency, and conventions are used to analyze all types of student writing. Because the six traits describe good writing, many teachers are helping students to look for these characteristics in their own writing and that of others. This provides students (and teachers)

a common vocabulary for communicating about and developing sound writing habits. (For more information on integrating analytical scoring into instruction see Spandel & Stiggins, 1989, and Spandel, 1992.)

Educators also use performance assessment of students for accountability purposes (i.e.,. to communicate to communities about student achievement), to inform building or district decision making (program evaluation, student certification), and for selection and placement decisions, either into advanced or remedial programs or into college.

Additionally, we have learned that performance assessment methodology holds the potential of helping us communicate to students and others what we value (Wiggins, 1988). If we value problem solving, cooperative learning, integrating writing and math across the curriculum, and critical thinking, our assessments must reflect this fact. Examples of attempts to use large-scale assessment to communicate valued outcomes are Connecticut's science performance assessments which require students to cooperate (Baron, 1990) and Vermont's inclusion of student dispositions (i.e.,. attitudes about the subject) in their scoring criteria for mathematics portfolios (Vermont Department of Education, 1991).

Targets. The kinds of student achievement targets being translated into performance assessments are those that require observation and considered professional judgment as the basis for evaluation. If we think of the various kinds of valued outcomes in education as being classified as (1) the mastery of subject matter knowledge, (2) the use of that knowledge to solve problems, (3) the exhibition of certain kinds of valued behaviors, (4) the creation of products that possess certain attributes and (5) the acquisition of certain affective responses, then performance assessments are primarily being used in the contexts of categories two, three and four.

For example, Oregon has developed and is pilot testing an analytical trait scoring scheme to evaluate mathematical problem solving for use in grades 3 and 8 as part of its state assessment. The four traits are conceptual understanding of the problem, procedural knowledge, problem solving skills and strategies, and communication[1]. The target is problem solving skill in mathematics and the ability to successfully communicate one's thinking.

Another example of the broad range of targets for student performance assessment is interactive speaking and listening--how well, for example, can students interact verbally with others in group discussions, social interactions, interviews, and instruction. The *English Language Skills Profile* (Hutchinson & Pollitt, 1987) has one exercise that involves a group discussion and another called a "paired interview." In the paired interview pupils are given written information about a proposed community project involving young people, and asked to discuss, in pairs, various aspects of its implementation with a view to decision making. There is a third person available to provide additional information upon request of the students. Students are assessed on their ability to interact appropriately with each other and the third person, appropriateness of comments, clarity of expression, willingness to cooperate, and the degree of support needed to complete the task.

Methods. The *methods* being used for student performance assessments vary in terms of the tasks that students perform, and in the criteria used to judge performance on the tasks.

Tasks include simulations, structured performance assessment tasks, portfolios and classroom exercises. The most common designs are direct observation of ongoing classroom events and the development and administration of structured performance assessment tasks. The former type of assessment tends to be informal, using checklists and rating forms developed without pilot testing. For example, the British Columbia Ministry of Education produced a document to assist classroom teachers with evaluation and instruction of oral communication (Jeroski, et. al, 1988). The handbooks contain a large number of checklists, observation forms, peer reviews, and self-reflection instruments for informal use in the classroom.

More structured performance assessment tasks are being designed in a number of contexts (see, for example, Baron, 1990; National Assessment of Educational Progress, 1987; California State Department of Education, 1989; Kanis, 1991; Larter, 1991; and Whetton, 1991). An example of a computer simulation is

presented by Shavelson, et. al (in press), in which the authors compare assessment of hands-on science laboratory tasks to several surrogates, such as lab notebooks, computer simulations and standardized test scores.

In addition to sound performance tasks, another key to effective student performance assessment is the careful development and application of proper criteria to use in the judgment process. Some assessments have criteria tied directly to the task, so that different criteria are developed for each task (e.g., California Department of Education, 1989 and Larter, 1991). Others, such as those involved in many writing assessments, find it more productive to develop broader criteria that can be applied across various tasks. For example, instead of looking for various features in a response (such as the presence of a graph), one would look for the ability of the student to employ appropriate solution methods, be flexible in the methods used, and switch methods when needed. The latter approach is more difficult to use because it requires that teachers and other users of the assessment completely understand what flexibility or efficiency really means and looks like across tasks.

One combination of tasks and criteria being given a great deal of attention in education these days is the use of the achievement portfolio. This application calls for the accumulation of examples of student work over time in a context where clear criteria have been established for the selection of entries into the portfolio, criteria have been developed for evaluating the work collected, and students play a key role in the assessment process by reflecting in a systematic manner on changes in their achievement as depicted by the work collected. One excellent example of this kind of assessment can be found in the work of Juneau Borough School District (Calkins, 1991). Each student portfolio includes several samples of student writing and reading collected at various times in the school year. Student progress is systematically rated using developmental continua. Students have input in deciding what will placed into their portfolios, and have the opportunity to explain why the pieces were selected for the portfolio and how he or she feels about him or herself as a reader and writer.

More information about performance assessment alternatives currently in use across the country is available from the Test Center at Northwest Regional Educational Laboratory[2], and in Arter (1989) and Arter & Spandel (1992).

Assessment of Teacher Performance

Purposes. The major reasons for conducting performance assessments of teachers are admission into teacher training programs, certification and licensure to teach, promoting professional development, accountability, and assuring minimum competence. For example, the California New Teacher Project has been exploring the use of performance assessments for teacher certification for several years. One set of prototypes (Murray, et. al, 1990) involved four simulations in which prospective teachers watched videotapes of typical language arts classroom situations and then answered a series of open-ended questions to assess their pedagogical knowledge. Performance criteria were developed to match each task. Stiggins and Duke (1989) drew a stark contrast between the uses of teacher performance assessment for professional development and accountability, depicting the key elements of assessments used for the former.

Targets. The aspects of teacher performance assessed include classroom management skills, instructional skills, and communication skills, among others. A wide variety of observational instruments and schemes for analyzing teacher's classroom behavior and products is collected in Good and Brophy (1987).

Methods. The performance assessment methods being used for teachers include classroom observation, portfolios, assessment centers, and simulations. For example, the *Teacher Assessment Project* at Stanford University tried portfolios and assessment center techniques (Teacher Assessment Project, 1988, 1989a, 1989b, 1989) to assess teacher subject area knowledge, pedagogical knowledge, and attitudes in biology and elementary literacy. For the literacy portfolio, teachers are asked to select four items that related to integrated language instruction, three that relate to creating a literate environment, and four about assessment of students. Teachers may also present an open entry and a reflective interpretation of any and

BEST COPY AVAILABLE

all entries. The related assessment center experience includes six exercises, some of which draw on the teacher's portfolio. Other exercises simulate teaching situations.

Assessment of Principal Performance

Purposes. The principle reasons for conducting performance assessments of principals are hiring, professional development, and accountability (formal job performance reviews).

The foremost example of performance assessment of principals for placement and professional development are the Assessment Centers operated under the auspices of the National Association of Secondary School Principals[3]. Over a three-day period at the center, principals are involved in six to eight exercises--leadership group exercises, in-basket exercises, fact-finding exercises, and structured interviews. Performance is observed by trained assessors who look for specific behaviors that are translated into scores in 12 areas--problem analysis, judgment, organizational ability, decisiveness, leadership, sensitivity, stress tolerance, oral communication, written communication, range of interest, personal motivation, and educational values.

Targets. The targets of performance assessment of principals are a variety of behaviors, styles, and skills. Management skills include such things as managing the budget, assessing student progress, and making the school run smoothly. Leadership involves vision setting, inspiring others to act, modeling the way, and inspiring the heart (Kouzes & Posner, 1988). Personality traits invalid such things as tolerance of ambiguity, sensitivity, and motivation. Instruments assessing styles have focused on such things as participatory leadership or directive leadership. Other knowledge and skills include ability to communicate orally and in writing and ability to solve problems. The twelve areas rated in the Assessment Center example cited above cover many of these dimensions.

Methods. Assessment tactics used for principal assessments have included structured interviews and on-the-job observation, as well as the assessment center and in-basket tasks described above. For example, the *Situational Leadership Instrument Package* (Hersey, et. al, 1982) includes an observational checklist (called the *Interaction Influence Analysis*) in which an observer keeps track of nine behaviors during an interaction between a leader and a subordinate.

Reviews of additional instruments for assessing the leadership and management qualities of school administrators (most of which are questionnaires and surveys) are available in Arter (1990).

Lessons Learned to Date

The Need For Clear Targets

The recent surge in interest in and development of performance assessments in education has brought benefits with it. For example, the obvious need to base subjective evaluations of sound criteria applied by carefully trained raters has necessitated the articulation of clear visions of the meaning of good performance. The 1980s was the decade for reexamining the valued outcomes of education. As that work has proceeded around the development of sound performance assessment criteria, we have acquired far clearer understandings of what it means, for example, to be an effective writer, reader, speaker, etc. This has tremendous potential for improving instruction as well as being essential for good performance assessment.

The Need For An Array of Assessment Tools

These sharper images of success have brought with them the realization that most of these valued outcomes are in fact far more complex than we had previously realized. This, in turn, has given new momentum to the drive for richer, more complete assessments of student achievement. Traditional paper and pencil tests, while still valuable tools, will never again be regarded as sufficient as a means of profiling student achievement. Rather, we now know that we must rely on a broad array of assessment tools to depict a broad array of valued outcomes.

The Need For Training

Our drive toward more diverse assessments has sensitized us to the need for new levels of assessment competence on the part of all educators and assessors. Sound performance assessments can only be developed and conducted by those who (a) possess a clear, highly-differentiated vision of the valued outcome, and (b) have mastered the craft knowledge needed to transform that target into appropriate performance exercises and performance criteria. Unfortunately, we have discovered that many charged with assessing student competence are not, in fact, qualified to do so.

Cost/Time/Technical Issues

Experimental application of performance assessment methodology in large-scale assessment contexts has revealed the great cost of this labor-intensive assessment alternative. These costs become prohibitive when considered in light of the lessons we are learning about the psychometric quality issues that must be addressed with performance assessments (Arter, 1991; Rothman, 1990; Valencia, 1989; Frechtling, 1991; Linn et. al, 1992). To meet accepted standards of validity (generalizability) and reliability (internal consistency), assessments often must include a variety of samples of student performance. If each sample carries with it high scoring costs, then the overall costs of an assessment that is sufficient in its breadth of exercises can be very high. Further, to meet accepted standards of reliability in the sense of objectivity or interrater agreement, very thorough rater training is essential and multiple judges are required to control for measurement error due to rater. All of this adds cost. In times of rapidly declining resources for education in general, rising assessment costs are a problem.

Issues Associated With High Stakes Testing

Many large-scale assessments are also high-stakes assessment--high school graduation, admission to college, report cards on schools, etc. These uses lead to their own problems, and indeed they are the same problems encountered previously with high-stakes testing--restricting curriculum, teaching the test (not just to the performance criteria), negative effects on students and teachers, the proliferation of a test-preparation industry that may or may really "work", and results that are, therefore, not valid. Just moving from multiple-choice tests to performance assessments will not solve these problems.

Performance Assessment As An Instructional Tool

This leaves educators on the horns of a dilemma. Many of the outcomes to be assessed are too complex to permit reliance on traditional paper and pencil objective tests. So we cannot return to yesteryear and rely solely on those. We must move forward and embrace performance assessment alternatives to create a complete profile of student achievement. But we cannot do that either, because the costs of such assessments is so astronomically high.

One possible solution to this dilemma might be to turn to teachers as the providers of the more complex student achievement information we desire. After all, they have the opportunity to gather that information needed over time, sampling with diverse exercises and providing the replications needed to produce valid assessment results. There are at least two problems with this plan. First, decades of neglect of teacher training in assessment has left teachers with neither the clear vision of achievement targets nor the performance assessment design expertise needed to play this critical role in the future of educational assessment.

Second, many performance assessments are designed in ways that tend to limit their usefulness instructionally, resulting in little incentive for teachers to *want* to put in the effort to gather high quality performance information.

Some features of current performance assessments which tend to limit instructional usefulness are performance criteria that are tied directly to individual tasks (so that the criteria change for each exercise);

holistic scoring; procedures that do not involve teachers in the scoring of performances; high stakes uses; and activities that do not make the students an interested partner. With respect to the latter point, tieing important decisions to the results of a performance assessment does not make the student an interested partner; an intimidated partner perhaps, but not one who is interested in an honest outside or self-appraisal of his or her status and progress.

An example of an instructionally useful performance assessment is the six-trait analytical procedure for writing used in Oregon. The same six traits describe good writing in general. Thus, the criteria are broad and not tied to particular exercises. This allows a consistent vocabulary for discussing writing across teachers and tasks. Students are made partners in the process by involving them in analyzing their own work and that of others using the criteria. Many teachers integrate the six traits with the writing process during peer review and revision as a consistent and powerful way to provide feedback using a common vocabulary. Others structure instruction around the traits so that, for example, students will spend some time thinking about and analyzing how organization can affect what an author is trying to say.

Teachers are made partners in this process by showing them how to use the model in instruction, and by involving as many teachers as possible in scoring statewide assessments. This procedure not only trains teachers in using the model, but also allows them to systematically apply them to large numbers of student papers, and to get a good idea of what student writing is really like at the various grade levels.

Articulating and applying performance criteria help teachers to know what "good" looks like and how students develop toward our goals for them. As Murphy and Smith (1990) state: "The benefits of portfolios lie as much in the discussions they generate among teachers--and among teachers and students--as in the wealth of information they provide." This is equally true of all good performance assessment because it forces us to articulate what we value in a performance and to apply it consistently to student work. Teachers and students learn in the process.

We would like to suggest that to have performance assessments that mean anything, we need to first ensure that teachers perceive them as good instructional tools and know how to use them as instructional tools. This will require a great deal of training. Thus, as we move through the early 1990s, we face a major unresolved performance assessment problem: We need teachers and want to take advantage of all they offer, but they simply are not equipped to do the job. Further, we appear not to have the resources with which to solve this immense national problem.

FOOTNOTES

1. For more information on this system, contact Michael Dalton, Oregon Department of Education, 700 Pringle Parkway S.E., Salem, OR 97310.

2. The Test Center at Northwest Lab has been collecting alternative assessment devices for several years. Annotated bibliographies of such instruments are available in the areas of reading, math, science and portfolios. Contact Judy Arter, Northwest Regional Educational Laboratory, 101 S.W. Main, Suite 500, Portland, OR 97204.

3. For more information about the NASSP Assessment Centers, write to the National Association of Secondary School Principals, 1904 Association Drive, Reston, VA 22091.

REFERENCES

Arter, J.A. (1989). Assessing communication competence in speaking and listening. Northwest Regional Educational Laboratory, 101 S.W. Main, Suite 500, Portland, OR 97204.

Arter, J.A. (1990). Assessing leadership and managerial behavior: A consumer guide. Northwest Regional Educational Laboratory, 101 S.W. Main, Suite 500, Portland, OR 97204.

Arter, J.A. (1991). Performance assessment: What's out there and how useful is it really? Paper presented at the annual meeting of the American Educational Research Association, Chicago.

Arter, J.A. & Spandel, V. (1992). NCME instructional module: using portfolios of student work in instruction and assessment, *Educational Measurement: Issues and Practice, 11*, pp. 36-44.

Berk, R.A. (1986). *Performance Assessment: Methods and Applications.* Baltimore MD: The Johns Hopkins University Press.

Calkins, A. (1991). Juneau integrated language arts portfolio for grade 1. Juneau Borough School District, 10014 Crazy Horse Drive, Juneau, AK 99801.

California State Department of Education (1989). A question of thinking. California State Department of Education, 721 Capitol Mall, P.O. Box 944272, Sacramento, CA 94244.

Fitzpatrick, R. & Morrison, E.J. (1971). Performance and product evaluation. In R.L. Thorndike (ed.), *Educational Measurement (2nd Ed.).* Washington DC: American Council on Education. pp. 237-270.

Frechtling, J.A. (1991). Performance assessment: Moonstruck or the real thing? *Educational Measurement: Issues and Practice, 10*, pp. 23-25.

Glaser, R. & Klaus, D.J. (1962). Proficiency measurement: Assessing human performance. In R.M. Gagne (ed.) *Psychological Principles in System Development*, New York: Holt, Rinehart & Winston, pp. 419-474.

Good, T.L. and Brophy, J. (1987). *Looking in Classrooms, 4th Ed.* New York: Harper & Row.

Hersey, P., Blanchard, K.H., Natemeyer, W.E., Hamilton, R.K., Keilty, J.W., Harrington, D., & Harrington, S.L. (1982). Situational leadership instrument package. University Associates, Inc., 8517 Production Ave., San Diego, CA 92121.

Hutchinson, C. & Pollitt, A. *The English Language Skills Profile*, London: Macmillan Educational.

Jeroski, S., Tolsma, C., Matheson, J., Clemens, A., Fisher, D., McIntosh, P., Olsen, D., Zwick, H., Derpak, D., Harding, R., Pietrasko, T., and Graham, S. *Enhancing and evaluating oral communication in the secondary grades: Teachers' resource package*, British Columbia Ministry of Education, Victoria, British Columbia, Canada.

Kanis, I.B. (1991). Ninth grade lab skills. *The Science Teacher*, January 1991, pp. 29-33.

Kouzes, J.M. and Posner, B.Z. (1988). *Leadership Practices Inventory*, University Associates, Inc., 8517 Production Ave., San Diego, CA 92121.

Larter, S. (1991). Benchmarks: The development of a new approach to student evaluation. Toronto Board of Education, 155 College Street, Toronto, Ontario, Canada M5T 1P6.

Linn, R.L., Baker, E.L., and Dunbar, S.B. (1992). Complex, performance-based assessment: Expectations and validation criteria. *Educational Researcher, 20*, pp. 15-23.

Murphy, S. and Smith, M.A. Talking about portfolios. *The Quarterly of the National Writing Project*, Spring 1990, pp. 1-3, 24-27.

Murray, S.L., Faddis, B.J., Sommer, A.T., Tyer, P., Calfee, R.C., Millman, J., Valencia, S. (1990). Assessment of pedagogical competence in elementary reading and language arts. California Department of Education,

National Assessment of Educational Progress (1987). Learning by doing: A manual for teaching and assessing higher-order thinking in science and mathematics. Report No. 17-HOS-80, Educational Testing Service, CN 6710, Princeton, NJ 08541.

Rothman, R. (1990). New tests based on performance raise questions. *Education Week, 10.*

Ryans, D.G. & Fredricksen, N. Performance tests of educational achievement. In E.F. Lindquist (ed.) (1951) Eduational Measurement. Washington DC: Amerian Council on Education, pp 455-493.

Shavelson, R.J., Baxter, G.P. & Pine, J. (in press). Performance assessment in science. *Applied Measurement in Education.*

Shulman, L. (1989). What makes a good teacher? *Teacher Magazine*, November 1989, pp. 35-36.

Spandel, V. (1992). *Integrating writing assessment and instruction in the classroom: Five day training institute for teachers.* Northwest Regional Educational Laboratory, 101 S.W. Main, Suite 500, Portland, OR 97204.

Spandel V., and Stiggins, R.J. (1989). *Creating writers: Linking assessment and writing instruction.* New York: Longman.

Stiggins, R.J. & Bridgeford, N.J. (1985) The ecology of classroom assessment. Journal of Educational Measurement, 22, 471-286.

Stiggins, R.J. & Duke, D.L. (1989) The Case for a Commitment to Teacher Growth. Albany NY: SUNY Press.

Teacher Assessment Project (1989a). Elementary literacy assessment center examiner's handbook. Stanford University, School of education, CERAS 507, Stanford, CA 94395.

Teacher Assessment Project (1989b). Biology examiner's assessment center handbook. Stanford University, School of education, CERAS 507, Stanford, CA 94395.

Teacher Assessment Project (1988). Portfolio development handbook for teachers of elementary literacy. Stanford University, School of education, CERAS 507, Stanford, CA 94395.

Valencia, S. (1989). Assessing reading and writing: Building a more complete picture. University of Washington, Seattle, WA 98195.

Vermont Department of Education (1991). *Looking beyond "the answer": The report of Vermont's mathematics portfolio assessment program.* Vermont Department of Education, 120 State Street, Montpelier, VT 05620.

Whetton, C., Ruddock, G., Hopkins, S., et. al (1991). Standard assessment tasks for Key Stage 1. National Foundations for Educational Research, The Mere, Upton Park, Slough, Berks S11 2Dq, England.

Wiggins, G. (1988). Rational numbers: Toward grading and scoring that help rather than harm learning. *American Educator*, Winter 1988, pp. 21-48.

Table 1
Applications of Performance Assessment in Education

	Students	Teachers	Principals
Purposes	Teacher instructional decision making Tool for instruction Student decision making Accountability Program evaluation Student certification, graduation, promotion College admission Communicate what is valued	Admission to training Certification Professional development Accountability	Hiring Professional development Accountability Admiss)ion to training
Targets	Subject matter knowledge Thinking processes Products, e.g., research reports Achievement related behavior, e.g. communication Affect, e.g., persistence, flexibility, self-confidence	Classroom management Instructional skills Subject matter knowledge Pedagogical knowledge Communication skills	Management skills Leadership behavior Personality traits/styles Problem solving Communication skills
Methods	Classroom observation Portfolios Structured performance assessment tasks Simulations	Classroom observation Assessment centers Simulations	On-the-job observation Assessment centers In-basket Structured interviews